DOCUMENT RESUME

ED 436 152 IR 019 777

AUTHOR	Koroghlanian, Carol M.; Sullivan, Howard J.
TITLE	The Use of Audio in Computer-Based Instruction.
PUB DATE	1999-02-00
NOTE	8p.; In: Proceedings of Selected Research and Development Papers Presented at the National Convention of the Association for Educational Communications and Technology [AECT] (21st, Houston, TX, February 10-14, 1999); see IR 019 753.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; *Audiovisual Instruction; *Auditory Stimuli; Comparative Analysis; *Computer Assisted Instruction; Computer Literacy; *Courseware; Design Preferences; Higher Education; *Instructional Design; *Instructional Effectiveness; Intermode Differences; Sex Differences; Student Attitudes; Tables (Data); Text Structure; Time on Task
IDENTIFIERS	*Text Density

ABSTRACT

This study investigated the effects of audio and text density on the achievement, time-in-program, and attitudes of 134 undergraduates. Data concerning the subjects' preexisting computer skills and experience, as well as demographic information, were also collected. The instruction in visual design principles was delivered by computer and included numerous illustrations. Subjects were randomly assigned to one of three presentation versions of the instruction: text only, full text-full audio, or lean text-full audio. No significant difference in achievement was found for the three treatment groups; however, significant differences in achievement were found for sex (females achieving more than males) and self-ratings of computer skill (subjects with higher self-ratings achieving more than those with lower self-ratings). While overall attitudes toward the instruction were favorable, subjects in the full text-full audio treatment responded less favorably than their counterparts in the other treatments. (Author/MES)

The Use of Audio in Computer-Based Instruction

Carol M. Koroghlanian
Arizona State University

Howard J. Sullivan
Arizona State University

Abstract

This study investigated the effects of audio and text density on the achievement, time-in-program, and attitudes of 134 undergraduates. Data concerning the subjects' preexisting computer skills and experience, as well as demographic information, were also collected. The instruction in visual design principles was delivered by computer and included numerous illustrations. Subjects were randomly assigned to one of three presentation versions of the instruction: Text Only, Full Text-Full Audio, or Lean Text-Full Audio. No significant difference in achievement was found for the three treatment groups; however, significant differences in achievement were found for sex (females achieving more than males) and self-ratings of computer skill (subjects with higher self-ratings achieving more than those with lower self-ratings). While overall attitudes towards the instruction were favorable, subjects in the Full Text-Full Audio treatment responded less favorably than their counterparts in the other treatments.

Introduction

Multimedia computer-based instruction (CBI) is increasingly used as an adjunct to instruction in schools and corporations and has typically incorporated text, graphics, and audio. Audio has often been added to CBI as an afterthought, to gain attention and increase motivation. When audio has been added as voiced material, it has often occurred as redundant reading of screen text. Little research has been performed to provide the instructional designer with guidelines for incorporating audio into CBI to promote learning. When guidelines do exist, such as those proposed by Barron (1995), they are frequently general and not always based on research. Indeed, little research exists to support the notion that adding audio to CBI can improve learning.

Most research on the use of audio in instruction has focussed on increasing learning through multichannel delivery and has traditionally investigated the mediums of television, videotape, 35-mm slide/tape, and audiocassette teamed with print. The underlying assumption has been that increasing the number of channels and the redundancy between channels will increase learning. This assumption stems from early research into multichannel communication. Hartman (1961) summarized and evaluated several studies that looked at audio-print and print presentation for instruction. He concluded that redundant audio-print instruction was more effective than either audio or print alone. This result was attributed to the cue summation theory that predicted learning would increase as the number of cues or stimuli are increased.

A competing theory to cue summation holds that only one sensory channel can access the brain at a time; implying that multichannel communication would not increase learning (Broadbent, 1958; Travers, 1964). In support of that view, a number of studies have found no advantage for audio-print over print alone in a variety of conditions with different age groups (Barton & Dwyer, 1987; Furnham, Gunter, & Green, 1990; Nugent, 1982; Van Mondfrans & Travers, 1964).

Most audio-text studies have looked at the use of text redundant audio where the audio is word for word the same as the text and both are presented simultaneously. A study by Barron and Kysilka (1993) and another by Barron and Atkins (1994) included text density as a treatment variable. Neither study found a difference in achievement between full text-no audio and lean text-full audio groups. If lean text with explanatory audio instruction is as effective as full text with no audio instruction, the implications for CBI design are considerable. Screen "real estate" is often at a premium—particularly when simulations or large labeled illustrations or graphics are important for the instruction. Some concepts are difficult to explain in words alone—pictures and illustrations may be necessary for clear understanding by the student.

Studies that have examined illustrations as well as text and audio are relatively scarce. Two studies (Mayer & Anderson, 1991; Mayer & Sims, 1994) looking at animations and narrations for teaching science concepts have found that animations accompanied by simultaneous narrations (audio) increase learning. Mayer and Anderson (1991) and Mayer and Sims (1994) believe that dual coding theory provides an explanation for how words and illustrations can be linked (Paivio, 1986). In Paivio's view, spoken and written language are both classified as verbal information. In dual coding, verbal information is selected and organized into verbal representations while visual information is selected and organized into visual representations. Finally, referential connections are built between the verbal and visual representations. This entire process must occur in short term or working memory.

Other researchers (Mousavi, Low, & Sweller, 1995; Tindall-Ford, Chandler, & Sweller, 1997) have indicated that when integration of diverse elements is crucial for learning (as is often the case for scientific or technical illustrations and accompanying text), audio-text instruction is effective and may decrease the burden on working memory that text only instruction might impose.

The present study examined the effects of audio and text density on achievement in a CBI with content that depended heavily on graphics, illustrations, and visual examples. Data concerning subject demographics, time in the CBI, attitude towards the CBI, and computer experience and skills were gathered.

The study was designed to investigate the following research questions:

- Does presentation version (text only, full text-full audio, and lean text-full audio) in a CBI affect student achievement, time-in-program and/or attitude?
- Does computer experience (years of use, frequency of use, and self-rated computer skills and confidence) affect student achievement and/or attitude in multimedia CBI?
- Does time-in-program affect student achievement and/or attitude?

Method

Subjects

One hundred thirty nine undergraduates (93 female, 46 male) enrolled in a computer literacy course at a large southwestern university participated in this study. The course, offered by the College of Education, fulfilled the university's undergraduate computer literacy requirement. While 35% of the subjects were education majors, the remainder represented a variety of majors ranging from Art to Zoology. The subjects were well distributed amongst freshman, sophomore, junior, and senior class standings (18%, 24%, 39%, and 18%).

Materials

A CBI, *Sending Your Message: Designing for Web and Print*, served as the instructional materials for the study. The CBI, developed by the first author, consisted of three topics: typography, design principles, and color. Each topic included information, examples, practice and feedback,.and review. The CBI was primarily linear in nature; however, the subjects could view previous screens within any of the three topics and could choose the sequence in which to view the topics. While the subjects had some sequence choices as they navigated the CBI, all subjects were required to view all instructional and review screens, and to complete all practice items before proceeding to the posttest.

The program was designed in three presentation versions corresponding to the three treatment conditions of the study. The first version was Text Only with some music and sound effects, but no voiced material, accompanying the instruction. The second version was Full Text-Full Audio with a male voice reading the instructional text that appeared on the screen word-for-word. It also contained the same music and sound effects as the Text Only version. The third version was Lean Text-Full Audio with the instructional text reduced to a bulleted outline accompanied by the audio from the Full Text-Full Audio version. Again, this version contained the same music and sound effects as the Text Only version. All three versions contained the same graphics, number of screens, practice items and reviews. All instructional screens included graphics to illustrate the concept or information presented in the text or audio. Both audio versions allowed the subjects to replay the audio.

Procedures

The subjects were randomly assigned to the three presentation versions. Students received instructions from the researcher, filled out a demographic/computer experience survey, worked through the CBI, took a posttest, and filled out an attitude survey. The two surveys were paper-based while the posttest was computer-based. All events occurred within the students' normally scheduled 110-minute laboratory period. All subjects wore headphones while working through the CBI.

Criterion Measures

Achievement was measured with a 30-item posttest administered by computer immediately after the instruction. The posttest consisted of multiple choice, matching and short answer items with each item worth 0 to 5 points depending upon the number of answers required by each item. Posttest items were similar in form and content to the practice items in the CBI. Posttest multiple choice and matching items were scored by computer, while short answer items were scored by the first author. The posttest KR-20 reliability was 0.80.

The attitude survey consisted of nine to eleven Likert scaled (five point scale from Strongly Agree to Strongly Disagree) items and two open-ended items regarding the usefulness, instructional quality and feelings towards the CBI. The number of Likert scaled items varied with the presentation version of the CBI. For the audio versions, two questions concerning the value of the audio were included. For the nine Likert scaled items common to all three CBI versions, the KR-20 reliability was 0.81.

Results

Achievement

Means and standard deviations for the posttest scores are shown in Table 1. The means for the three versions were very similar and the results of an ANOVA run on the posttest means indicated that there was no significant difference in scores for the different CBI versions, $F(2, 131) = 0.725$, $p > 0.05$.

Table 1. Mean Posttest Scores and Standard Deviations by CBI Version.

Version	N	Mean Score* (M)	Standard Deviation (SD)
Text Only	47	47.23	6.99
Full Text-Full Audio	40	45.75	7.10
Lean Text-Full Audio	47	47.30	5.95
All Versions Combined	134	46.81	6.66

*A maximum score of 63 points was possible on the posttest.

Time-in-Program

Means and standard deviations for the time-in-program are shown in Table 2. A one-way ANOVA indicated that there was a significant difference in the means, $F(2, 131) = 18.50$, $p < 0.05$. A post hoc Bonferroni test identified the source of the difference. The mean for the Text-Only version was significantly different than the means for the Full Text-Full Audio and the Lean Text-Full versions at the $p < 0.05$ level. There was no significant difference for time-in-program between the Full Text-Full Audio and Lean Text-Full Audio versions.

Table 2. Mean Time-in-Program and Standard Deviations by CBI Version.

Version	N	Mean Time-in-Program (seconds) (M)	Standard Deviation (SD)
Text Only	47	1830	596
Full Text-Full Audio	40	2523	712
Lean Text-Full Audio	47	2515	574
All Versions Combined	134	2277	703

There was no correlation between time-in-program and achievement, $F(1, 132) = 0.447$, $p > 0.05$ nor were any correlation's between time-in-program and attitude discovered.

Attitudes

The means and standard deviations for the attitude responses are shown in Table 3. A MANOVA for attitude scores indicated no significant difference for the different versions, $F(9, 246) = 1.38$, $p > 0.05$. Follow-up univariate tests were performed for individual items.

Table 3. Mean Attitude Scores and Standard Deviations by CBI Version and Versions Combined.

Item		Text Only	Full Text-Full Audio	Lean Text-Full Audio	All Versions Combined
1. I liked this program.	Mean SD	3.81 1.04	3.50 0.88	3.83 0.80	3.72 0.92
2. I would recommend this program to other students.	Mean SD	3.72 0.97	3.58 0.96	3.87 0.72	3.73 0.89
3. I preferred learning about design in a computer program rather than in a lecture format.	Mean SD	4.11 1.15	3.65 1.25	3.85 1.09	3.88 1.17
4. This was a good way to learn about design.	Mean SD	4.17 0.99	3.90 0.98	4.13 0.86	4.07 0.94
5. I learned a lot about design.	Mean SD	3.89 0.98	3.50 0.78	3.87 0.93	3.77 0.92
6. I plan to use what I have learned in future work.	Mean SD	3.85 1.02	3.75 0.74	4.22 0.66	3.95 0.85
7. This program was easy.	Mean SD	4.15 0.86	4.03 0.86	4.11 0.71	4.10 0.81
8. I would like to learn more about design.	Mean SD	3.68 1.11	3.48 0.91	3.70 0.81	3.62 0.95
9. I tried hard to do well in this program.	Mean SD	3.66 1.01	3.20 0.97	3.72 0.91	3.54 0.98
10. The audio helped me learn better.	Mean SD		3.18 1.15	3.61 1.14	3.41 1.16
11. I liked having the audio.	Mean SD		3.23 1.10	3.78 1.13	3.52 1.15
12. The amount of practice was	Mean SD	1.11 0.38	1.30 0.56	1.51 0.82	1.30 0.63
13. The amount of text on the screens was	Mean SD	1.19 0.45	1.23 0.42	1.15 0.47	1.19 0.45
Number of Subjects (N)		47	40	47	134

Notes: Items 1-9 were measured on a five-point scale from 1 to 5 (Strongly Disagree to Strongly Agree).
Items 10-11 were not administered for the Text Only version.
Items 12-13 were measured on a three point scale from 1 to 3 (About Right, Too Much, Too Little)

Univariate analyses revealed that there were significant differences for four items among the different versions with subjects in the Full Text-Full Audio version responding more negatively to three of these items than their counterparts in another version. Subjects in the Full Text-Full Audio version (M = 3.75, SD = 0.74) responded more negatively than those in the Lean Text-Full Audio version (M = 4.22, SD = 0.66) to the statement, "I plan to use what I have learned in future work", $F(2, 130) = 3.89$, $p < 0.05$. A similar pattern was observed for the statement, "I tried hard to do well in this program", with subjects in the Full Text-Full Audio version (M = 3.20, SD = 0.97) responding more negatively than those in the Lean Text-Full Audio version (M = 3.72, SD = 0.91), $F(2, 130) = 3.64$, $p < 0.05$. In terms of the amount of practice, subjects in the Lean Text-Full Audio version (M = 1.51, SD = 0.82) were less inclined to respond that "The amount of practice was about right" than subjects in the Text Only version (M = 1.11, SD = 0.38), $F(2, 129) = 5.049$, $p < 0.05$. Finally, subjects in the Full Text-Full Audio version (M = 3.23, SD = 1.10) responded more negatively than those in the Lean Text-Full Audio version (M = 3.78, SD = 1.13) to the statement, "I liked having the audio", $F(1, 84) = 5.33$, $p < 0.05$.

Overall, the subjects in the Full Text-Full Audio version appeared to respond less favorably to the CBI if responses to items one to nine in Table 3 were averaged (Text Only M = 3.89, Full Text-Full Audio M = 3.62, Lean Text-Full Audio M = 3.92).

Additional Findings

Information about the subjects, demographics, as well as computer experience, was collected prior to starting the CBI. Several questions regarding computer experience were asked including: a self-rating of computer skills, confidence in succeeding at computer tasks, hours per week of computer use, and years of computer use. These variables and two demographic variables (Age and Sex) were examined in conjunction with achievement.

Sex ($r = -0.229$, $p < 0.05$), self-rating of computer skills ($r = 0.252$, $p < 0.05$), hours per week of computer use ($r = 0.186$, $p < 0.05$), years of computer use ($r = 0.210$, $p < 0.05$), and self-rating of confidence in succeeding at computer tasks all showed some correlation with achievement. Not surprisingly, self-rating of computer skills was moderately correlated to confidence, hours per week of computer use, and years of computer use. In summary, females achieved significantly more than males and subjects with higher self-rating of computer skills achieved significantly more than those with lower self-ratings of computer skills.

The status variables were also examined in conjunction with the attitude items. There was a significant correlation between sex and attitudes toward amount of practice with females being more inclined towards stating that there was enough practice, $F(1,130) = 5.732, p < 0.05$ than males. Sex also correlated with the statement, "I tried hard to do well in this program", $F(1,131) = 5.570, p < 0.05$ with females indicating more effort than males. The more confident subjects felt when dealing with computers, the more likely they were to agree with the statement, "This program was easy", $F(1,131) = 6.060, p < 0.05$. Age correlated with the statement, "I would like to learn more about design", $F(1,131) = 5.297, p < 0.05$ with older subjects more apt to rate the item higher than younger subjects. For the audio versions, self-rating of computer skills correlated with the statement, "The audio helped me learn better", $F(1,84) = 6.277, p < 0.05$ with more skilled subjects responding more favorably to the audio.

Discussion

The present study examined the effects of varying levels of audio and text density on achievement, time-in-program, and attitudes in a CBI. The computer experience, skills, and demographics of the subjects were also examined. The CBI was designed following instructional design principles and the content required extensive use of graphics for illustrations and visual examples. All instructional screens contained some graphic material that illustrated the concept presented in the text or audio.

Achievement

The CBI was fairly successful at teaching the subjects about basic design, a topic which few had any preexisting knowledge. While the majority of the instructional objectives were knowledge and comprehension objectives, a few were analyze and evaluate. The number of analyze and evaluate items on the posttest were too few to allow for interpretation in a broader context but subjects appeared to score about as well on those as on the knowledge and comprehension items.

There were no significant differences in achievement based on the presentation version of the CBI. Text Only was as effective as Full Text-Full Audio, a result that echoes many other studies (Barron & Kysilka, 1993; Barton & Dwyer, 1987; Furnham, Gunter, and Green, 1990; Nugent, 1982; Rehaag & Szabo, 1995). Again similar to other studies (Barron & Aikens, 1994; Barron & Kysilka, 1993) employing CBI, the lean text version was as effective as a full text version. By shifting part of the verbal portion of instruction to the audio channel, more screen "real estate" may be devoted to detailed graphics, illustrations, visual examples and simulations.

The present results do not appear to provide evidence for the cue summation theory, as there was no difference in achievement between the treatments. Some might argue that the present results provide support for the theory that only one sensory channel has access to the brain at one time (Broadbent, 1958; Travers, 1964); thus predicting no difference in achievement for the three presentation versions. Dual coding theory would view the text and audio as equivalent (verbal information) and processed simultaneously with the graphics (visual information). In this view, the three treatments are equivalent and similar results would be expected for each. Research examining graphics, audio and text (Mayer and Anderson, 1991; Mayer & Sims, 1994; Tindall-Ford, Chandler, & Sweller, 1997) might have predicted that the Lean Text-Full Audio version would do better than the Text Only version and possibly better than the Full Text-Full Audio version.

Time-in-Program

Not unexpectedly, the Text Only version took much less instructional time than either of the audio versions—amounting to a difference of about 10 minutes (total instructional time was 30 minutes for the Text Only version and 40 minutes for the Audio versions). While subjects could have advanced through the program without waiting for the audio to finish, the vast majority of subjects waited for the completion of the audio segment before continuing to the next screen. As time-in-program was more likely a function of the presence or absence of audio, it is not surprising no correlation between time and achievement was discovered. Interestingly, time-in-program was not correlated with attitude. It might have been expected that longer instructional time might have engendered some degree of negative attitude but such was not the case. The present study clearly indicates that significant amounts of audio in CBI will substantially increase the amount of time required for instruction.

Attitudes

While subjects' general attitudes were favorable ($M = 3.81$ averaged over all subjects, all treatments for items 1-9 in Table 3), subjects in the Full Text-Full Audio version were less enthused than their counterparts in the other versions (Text Only $M = 3.89$, Full Text-Full Audio $M = 3.62$, Lean Text-Full Audio $M = 3.92$). In fact, the mean response for the Full Text-Full Audio version was the least positive of the three treatments on each of the first nine attitude items. The means for the Text Only and Lean Text-Full Audio versions were quite similar for seven of the nine items with the Lean Text-Full Audio usually being the more positive of the two.

Additional Findings

Implications

The present study could have implications for the instructional design of CBI where screen "real estate" is needed for something other than instructional text. This is particularly true for simulations and concepts difficult to explain with text alone. Science and technical material as well as visual arts and design are particularly well served by detailed illustrations, visual examples, animations and simulations. If the screen can contain a graphic or labeled illustration with a minimum of explanatory text while audio supplies the detailed explanation, a major design problem can be overcome.

The current findings also provide support for avoiding redundant audio in most situations as an unnecessary expense that adds considerable instructional time for no apparent learning gain. For some audiences, redundant audio may even decrease positive attitudes towards the instruction. Redundant audio should probably be reserved for situations where the audience includes poor readers or used in brief segments to emphasize critical points or issue warnings.

Generalizability

Caution should be exercised before generalizing results from this study. While the present findings support previous research concerning the effectiveness of Lean Text-Full Audio (Barron & Atkins, 1994; Barron & Kysilka, 1993), much is left unresolved. Both of the Barron studies involved technical material with mostly comprehension and knowledge objectives. The present study involved non-technical material and included a few analyze and evaluate objectives. Three studies with two content areas do not constitute much evidence. More research needs to be done on Lean Text-Full Audio in conjunction with higher level objectives and diverse content areas before guidelines for instructional designers can be promulgated with any certainty. Shifting instructional text to audio may only be effective for certain types of content and objectives. Combinations of content and objectives need to be explored in greater detail.

References

Barron, A. E. (1995). Digital audio in multimedia. Educational Media International, 32 (4), 190-193.

Barron, A. E. & Atkins, D. (1994). Audio instruction in multimedia education: Is textual redundancy important? Journal of Educational Multimedia and Hypermedia, 3 (3-4), 295-306.

Barron, A. E. & Kysilka, M. L. (1993). The effectiveness of digital audio in computer-based training. Journal of Research on Computing in Education, 25, 277-289.

Barton, E. A. & Dwyer, F. M. (1987). The effect of audio redundancy on students' ability to profit from printed-verbal/visualized instruction. International Journal of Instructional Media, 14, 93-98.

Broadbent, D. E. (1958). Perception and communication. New York: Pergamon Press.

Furnham, A., Gunter, B., & Green, A. (1988). Remembering science: the recall of factual information as a function of presentation mode. Applied Cognitive Psychology, 4, 203-12.

Hartman, F. R. (1961). Single and multiple channel communication: A review of research and a proposed model. AV Communication Review, 9, 235-262.

Mayer, R. E. & Anderson, R. B. (1991). Animations need narrations: an experimental test of a dual coding hypothesis. Journal of Educational Psychology, 83, 484-90.

Mayer, R. E. & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. Journal of Educational Psychology, 86 (3), 389-401.

Mousavi, S., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. Journal of Educational Psychology, 87, 319-334.

Nugent, G. C. (1982). Pictures, audio, and print: Symbolic representation and effect on learning. Educational Communication Technology Journal, 61, 179-211.

Paivio, A. (1986). Mental representations: A dual coding approach. Oxford, England: Oxford University Press.

Rehaag, D. M. & Szabo, M. (1995, February). An experiment on effects of redundant audio in computer based instruction on achievement, attitude, and learning time in 10th grade math. Paper presented at the Annual Meeting of the Association for Educational Communications and Technology, Anaheim, CA. (ERIC Document Reproduction Service No. 380 123).

Tindall-Ford, S., Chandler, P., Sweller, J. (1997). When two sensory modes are better than one. Journal of Experimental Psychology Applied, 3, 257-287.

Travers, R. M. W. (1964). The transmission of information to human receivers. AV Communication Review, 12 (6), 373-385.

Van Mondfrans, A. P. & Travers, R. M. W. (1964). Learning redundant material presented through two sensory modalities. Perceptual and Motor Skills, 19, 743-751.

NOTICE

REPRODUCTION BASIS

☒ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").